REMARKS

The amendments and remarks presented herein are in response to the Office Action dated September 24, 2002.

Claims 1-35 are pending.

The Examiner stated that the title of the invention is not descriptive. Applicants have amended the title to be more descriptive of the invention.

The Examiner reminded Applicants to update the cited co-pending application serial numbers. Applicants have amended the specification at paragraph 0001 to properly cite the relevant serial numbers.

The Examiner rejected claims 1-35 under 35 U.S.C. §102(b) as being anticipated by Damashek, U.S. Patent No. 5,418,951. This rejection is respectfully traversed.

On January 22, 2003, the Examiner and the Applicants' representative conducted a telephone interview to discuss the pending Office Action. Applicants thank the Examiner for the opportunity to discuss the case in such a manner. In the course of the interview, the Examiner and the Applicants' representative discussed various distinctions between the invention claimed herein and the cited art. The following remarks summarize the discussion.

Claim 1 recites:

> A computer-implemented method of text equivalencing from a string of characters comprising:
> modifying the string of characters using a predetermined set of heuristics;

comparing the modified string with a known string of characters in order to locate a match;
responsive to not finding a match, forming a plurality of sub-strings of characters from the string of characters; and
using an information retrieval technique on the sub-strings of characters to determine a known string of characters equivalent to the string of characters.

The claimed invention is directed to a method for locating a character string that is equivalent to a query string. The invention attempts to locate a match for the query string by modifying the string and then comparing the string with a known string. If no match is found, the invention forms sub-strings of characters and attempts to determine an equivalent known string.

By contrast, Damashek does not perform character string matching. Damashek merely uses n-gram distributions to identify, retrieve, and sort documents according to language or topic. More particularly, Damashek determines the frequency distribution of n-grams within a document, and compares the frequency distribution with representative documents in order to categorize according to language or topic. Nowhere in Damashek is there any teaching or suggestion of finding a match for a character string, or attempting to find a string that is equivalent to a query string, as claimed herein.

For example, claim 1 recites "comparing the modified string with a known string of characters in order to locate a match." In this step, the claimed invention determines whether the modified string matches a known string. The Examiner stated that this step is disclosed at col. 4, lines 56-59 of Damashek. However, the cited portion of Damashek merely discusses comparing the *frequency of occur-*

rence of n-grams in the unidentified text with the frequency of occurrence of n-grams in the text of known languages. Under the assumption that documents in a given language will have similar n-gram distribution frequencies, Damashek attempts to determine the language of the unidentified text based on the results of the comparison of frequencies of occurrence. (See Damashek, col. 4, lines 24-27). Damashek compares *frequencies of occurrence*, and does not teach or suggest any technique for comparing *character strings* in order to locate a match or determine a text equivalent, as claimed herein.

Claim 1 further recites "responsive to not finding a match, forming a plurality of sub-strings of characters from the string of characters." In this step, the claimed invention forms sub-strings that are subsequently used in locating an equivalent string. The Examiner stated that this step is disclosed at col. 3, lines 21-24 and col. 4, lines 24-27 of Damashek. However, col. 3, line 21-24 of Damashek merely discuss the use of a pattern recognition technique based on n-gram comparisons; col. 4, lines 24-27 describe the hypothesis that documents that are similar in language and/or topic tend to contain similar n-gram distributions. Neither of these cited portions mentions any technique of forming sub-strings as claimed herein. Furthermore, although Damashek does disclose parsing text into n-grams (col. 4, line 50) as part of a technique for categorizing documents according to language or topic, Damashek does not form sub-strings *in response to not finding a match*, as specifically recited in claim 1.

Finally, claim 1 also recites "using an information retrieval technique on the sub-strings of characters to determine a known string of characters equivalent to the string of characters." As discussed above, in this step the claimed invention attempts to determine string equivalents. The Examiner stated that this step is disclosed at col. 9, lines 64-66 of Damashek. However, the cited portion of Damashek merely notes that a query concerning a topic of interest will result in documents retrieved on that topic that are written in the language of the query. This is to clarify that, in categorizing or retrieving documents according to topic, documents written in different languages will generally use different n-grams to represent the same topic, and therefore will not usually be placed in the same category. The discussion of this limitation of Damashek is entirely unrelated to the claimed step of determining string equivalents using an information retrieval technique.

Claim 12 is a system claim reciting elements corresponding to the method steps of claim 1. Claim 23 is a computer-readable medium claim comprising computer-readable code for performing the steps recited in claim 1. Claims 2-11, 13-22, and 24-35 are dependent claims that incorporate the limitations of claims 1, 12, or 23. The above arguments therefore apply to claims 2-35.

Accordingly, on the basis of the above remarks, consideration of this application and the early allowance of all claims herein are requested.

Should the Examiner wish to discuss the above remarks, or if the Examiner believes that for any reason direct contact with Applicants' representative would help to advance the prosecution of this case to finality, the Examiner is invited to telephone the undersigned at the number given below.

Respectfully submitted,
Ted E. Dunning and
Bradley D. Kindig

Dated: JAN 24/03

By: _____
Amir H. Raubvogel,
Reg. No. 37,070
Fenwick & West LLP
801 California Street
Mountain View, CA 94041
(650) 335-7276

Attachment: Version with Markings to Show Changes Made

VERSION WITH MARKINGS TO SHOW CHANGES MADE

(AMENDMENT A)

In the specification:

For the convenience of the Examiner, paragraph 0001 is reproduced below:

The present application claims priority from U.S. Patent Application Serial No. [_____] 09/846,823 for "Relationship Discovery Engine," filed April 30, 2001, the disclosure of which is incorporated herein by reference. The present application also claims priority from provisional U.S. Patent Application Serial No. 60/201,622, for "Recommendation Engine," filed May 3, 2000, the disclosure of which is incorporated herein by reference.